OPTIMAL RESOURCE ALLOCATION TO MAXIMIZE MODEL ACCURACY

BY

GUILLERMO LEO DANKERT

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Civil and Environmental Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Adviser:

ABSTRACT

In the modelling of civil engineering structures, the engineer should make several assumptions to obtain a reasonable prediction of the behavior of the structure. Experience in projects in the professional practice shows that model assumptions made mechanically, without getting into specific details, lead to inaccurate models and to expensive resources allocation that does not render the optimal results.

An example within civil engineering is the modelling of a structure, where – typically - structural engineers model the structure above the foundation and geotechnical engineers model the foundation and the soil behavior. These two areas have different error tolerances, mainly because the mechanical uncertainty of manufactured materials as concrete and steel is smaller than the uncertainty of soil. If these differences are not addressed correctly, the optimization process to get more accurate results is done incorrectly. This kind of situation is also common to other interaction with seismic engineering or hydrological engineering, among others.

To address the limitations, it is necessary to consider the problem as a system, rather than dealing with the local behavior of the different parts.

This work proposes a procedure based on the importance of random and categorical variables to address the different sources of uncertainty. Considering individual component models as categorical variables, the procedure follows a multi-model uncertainty propagation approach. The developed procedure is applied to an example of a cantilever beam on clay foundation.

Keywords: Model uncertainty, Structural engineering, Soil structure interaction, multi-model uncertainty propagation

ACKOWLEDGMENTS

I would first like to thank my thesis advisor Dr. Paolo Gardoni of the School of civil and environmental engineering at the University of Illinois at Urbana Champaign. His tidiness and intellectual clarity demanded a clear and rigorously mathematic work. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever he thought I needed it.

I would also like to thank Dr. Scott Olson who was involved in this research project, assisting with soil mechanics definitions and expertise. Without his passionate participation and input, this work could not have been successfully conducted.

My appreciations are extended to my team colleagues, Armin Tabandeh and Hao Xu whose uncountable hours in explaining steps of the application example were required to complete this work. and dedication in reviewing my work were a milestone in this path. I would also like to acknowledge Roberto Guidotti and Alvaro Canga for the dedication in reviewing my work. I am gratefully indebted to theirs very valuable comments on this thesis.

Finally, I must express my very profound gratitude to my spouse Maria de los Angeles for providing me with unfailing support, continuous encouragement and immense patience throughout my years of study and through the process of researching and writing this thesis. Thank you.

TABLE OF CONTENTS

CHAPTER 1 - INTRODUCTION AND PROBLEM STATEMENT

In order to implement reliability based design in engineering within a probabilistic framework, the assessment methodology should explicitly account for the prevailing uncertainties in the model. Following Gardoni et al (2002) we define the model as the mathematical expression relating one or more quantities of interest $\boldsymbol{y} = (y_1, y_2, \dots)$ to a set of measurable variables $\boldsymbol{x} = (x_1, x_2, \dots)$. This mathematical model is a combination of parametrized probabilistic sub-models $f_x(\boldsymbol{x}, \Theta_f)$ describing the distribution of the random vector $\boldsymbol{x}$, and a set of parameterized physical sub-models $\boldsymbol{y} = g_i(\boldsymbol{x}, \Theta_g),\ i = 1,2, \dots, m$ describing relations between the quantities $\boldsymbol{x}$ and derived quantities $\boldsymbol{y}$. The sub-models $f_x(\boldsymbol{x}, \Theta_f)$ and $g_i(\boldsymbol{x}, \Theta_g)$ are invariably imperfect mathematical idealizations of reality and contain uncertain errors. Their parameters, Θ_f and Θ_g, are usually assessed through a process of ‘‘fitting” these sub-models to observed data.

In general, alternative candidates for each of these models may be considered, so the art of modelling requires the definition of each of the physical and probabilistic sub-models. This decision is typically non-trivial.

To fit the final model, we would require data from the behavior of similar projects, which in large civil engineering constructions is not easily found, and even in that case it is very difficult to get reliable data. Therefore, the solution of fitting the complete behavior is not achievable. What is done instead is data collection to calibrate sub-model behavior (soil test, concrete test, flow measures, …) demanded by the modeler to reduce the total uncertainty.

Furthermore, the most common physical sub-models $g_i(\boldsymbol{x}, \Theta_g)$ in high complexity civil engineering problems are finite element (FE) models. These models show high precision in solving deterministic problems but are not suitable (in the commercial area) for a proper reliability analysis. FE models can have a high computational burden; considering the probability of failure expected for civil engineering projects (10^{-2} to 10^{-7}), this make a combination that requires special attention, this topic is in discussed in 2.2.

In common practice the component models selection is done recurring to previous experience (from others or own), nevertheless it is very difficult for the modeler to have experience in all the areas of the problem, in order to select the most accurate sub-models. Two main mistakes are regularly done, the first one is the selection of only one component model with no further justification, and the second is that the sub-model selection is done without considering the precision of the other component models, obtaining consequently a combination of different sub-models with randomly selected precisions.

Modelling in applied science and engineering involves the making of decisions (Field Jr, et al., 2007). Then the decision for model selection allows us to quantify the consequences of choosing one model of the system over another. The consequences considered in this work are the variation that model selection produces on the reliability of the structure.

The objective of this work is to develop a methodology capable of accounting for the different sources of uncertainty and perform a fair comparison capable of deciding where to allocate the resources available by the modeler to increase the accuracy of the system. This assessment of component model selection is done in the design phase of the problem, i.e. when data of the complete model behavior is not available.

CHAPTER 2 - LITERATURE REVIEW

There are several publications about uncertainty quantification, a small review of the relevant topics for this work as well some introduction to uncertainty through FEM and surrogate models is done to give the reader who is not familiar with these topics some references.

2.1 Types of uncertainties

Following the definition for sources of uncertainty provided in (Der Kiureghian, et al., 2009), we identify:

1. Uncertainty inherent in the basic random variables $\boldsymbol{x}$, such as the uncertainty inherent in material property constants and load values, which can be directly measured. For example, the compressive strength of concrete.

2. Uncertain model error resulting from selection of the form of the probabilistic sub-model $f_x(\boldsymbol{x}, \Theta_f)$) used to describe the distribution of basic variables. For example, the results of adopting a normal distribution are not the same as the ones obtained by adopting a beta distribution for a certain variable x.

3. Uncertain modelling errors resulting from selection of the physical sub-models $g_i(\boldsymbol{x}, \Theta_g)$, used to describe the derived variables. For example, assuming linear behavior for the mechanical behavior of some material or using a non-linear approach.

4. Statistical uncertainty in the estimation of the parametersΘ_f of the probabilistic sub-model.

5. Statistical uncertainty in the estimation of the parameters Θ_g of the physical sub-models.

6. Uncertain errors involved in measuring of observations, based on which the parameters Θ_f and Θ_g are estimated. These include errors involved in indirect measurement, e.g., the measurement of a quantity through a proxy, as in non-destructive testing of material strength.

7. Uncertainty modeled by the random variables Y corresponding to the derived variables y, which may include, in addition to all the above uncertainties, uncertain errors resulting from computational errors, numerical approximations or truncations. For example, the computation of load effects in a nonlinear structure by a finite element procedure employs iterative calculations, which invariably involve convergence tolerances and truncation errors.

Reliability analysis design considers uncertainties of the type 1 as its input. Furthermore, importance measures asses the relevance that each of the x variable randomness has in the probability of failure (Der Kiureghian, et al., 1985).

A methodology for reducing error of the type 4, 5 and 6 is developed in Gardoni et al (2002 and 2003), and it is applied as an example in reinforced concrete column design under seismic load, as well as other examples over bridge columns (Huang, et al., 2010) and applied to columns reinforced with FRP (Tabandeh, et al., 2014) (Tabandeh, et al., 2015).

Type 7 error is not going to be addressed directly in this work, but precision in the order of 10^{-10}, is going to be used when using FEM, which is despicable compared to other sources of error.

An approach of solving type 2 and 3 errors is done in Sun et al (unpublished) where model selection is done to better fit a data set to a selection of candidate models, but this approach does not consider the possibility of design where data is unavailable.

Then the objective of this work is to develop a methodology capable of accounting for the uncertainty of the first 6 cases, and do a fair comparison among them to be able to decide where to allocate the resources available by the model in order to increase the accuracy of the system.

2.2 Uncertainty propagation through finite element model (FEM)

Finite element modelling is probably the most common way of solving complex engineering problems, its capabilities cover most of the existent problems in current practice and are still being increased in the research area. However, the method is deterministic, that means that is not possible to contemplate uncertainty in the input parameters. A powerful tool in computational stochastic mechanics is the stochastic finite element method (SFEM). SFEM is an extension of the classical deterministic FE approach to the stochastic framework i.e. to the solution of stochastic (static and dynamic) problems involving finite elements whose properties are random. From a mathematical point of view, SFEM can be seen as a powerful tool for the solution of stochastic partial differential equations (PDEs) and it is treated as such in numerous studies where convergence and error estimation issues are examined in detail. In fact, these two aspects of SFEM are complementary and inter-dependent. The considerable attention that SFEM received over the last decade can be mainly attributed to the spectacular growth of computational power rendering possible the efficient treatment of large-scale problems (Stefanou, 2009).

A complete review of SFEM is done in Stefanou (2009). We can reduce the SFEM to two options: intrusive SFEM and non-intrusive SFEM.

In intrusive SFEM, the randomness is considered in the element formulation. The SFEM comprises three basic steps: the discretization of the stochastic fields representing the uncertain system properties, the formulation of the stochastic matrix (first at the element and then at the global-system level) and finally, the response variability calculation (Stefanou, 2009).

In the non-intrusive approaches, a surrogate model is created through a response surface without interfering with the FE procedure i.e. without directly modifying the element matrix. This is why these methods can take advantage of powerful deterministic FE codes and use them as a black-box.

Two alternatives to create surrogate model where analyzed, physics based response surfaces and polynomial chaos expansion (PCE) based surfaces. Physics based surrogate models is an application of the work done by Gardoni et al (2002 and 2003). A very complex FEM is set as the data to be fitted and physics based functions together with simple explanatory functions are fitted through a Bayesian methodology. Further explanation of the method is going to be done in 2.3. Surrogate models for reliability analysis . By PCE based surfaces a polynomial family response function is built as a surrogate model, the reader is referred to Sudret et al (2002) and Marelli et al (2015).

Even though the two approaches are mathematically equivalent, the physics based approach has the advantage that all parameters have a physical meaning, making the final model easier to understand from an engineering approach. Regarding that these models are used to take high risk decision, the comprehensiveness of the model is a desirable quality.

The assessment of the uncertainty propagation through FEM requires: (a) Reliability software allowing to determine the design point for FORM analysis and to perform simulation, importance sampling, etc. and (b) Deterministic finite element model.

Several reliability software programs are currently available, both in the research area, open source, and in the commercial area. A full analysis of available reliability software until 2005 was done by Pellisetti et al (2006). Both in the professional area and the structural design research area, is not common to be familiar with several programing languages. The most common language is the MATLAB code. Three MATLAB based programs are available for free: *FERUM* (Der Kiureghian, et al., 2006) (Haukaas, et al., 2005), *UQLAB* (Marelli, et al., 2015) and *COSSAN* (Patelli, et al., 2012). For the purpose of this project, FERUM has been used.

2.3 Surrogate models for reliability analysis

The goal of the design is set to solve reliability problems that involve the structural capacities at the component level, $C = (C_1, C_2, \dots)$, and the corresponding demands, $D = (D_1, D_2, \dots)$ .Component capacities are defined as the set of forces and deformations that a component can carry without failing, e.g. the maximum shear force or deformation that a column can sustain. The component demands are defined as the forces and deformations to which a component is subjected to for a given system demand, e.g. an earthquake ground motion, extreme wind loads, etc. The probability of failure of a component is then defined as the probability that the demand D_k is greater than or equal to the corresponding capacity C_k, where k ranges over all the possible modes of failure, e.g. failure in shear or excessive deformation. The failure of the structural system is defined in terms of the component failure events, and the corresponding probability is computed by use of the methods of structural system reliability.

Predictive capacity and demand models in current structural engineering practice are typically deterministic and often on the conservative side. The objective of building surrogate models is to get a fast model capable of predicting correctly the behavior of a complex model.

Following the approach of the referenced work rather than developing new models, correction terms to existing explicit or simple procedure deterministic capacity and demand models were developed through the use of a set of ‘’explanatory’’ functions. With these functions we are able to identify terms that correct the bias in the surrogate explicit model.

The construction of the surrogate model needs:

- An existing deterministic capacity (or demand) model,
- A complex FEM capable of faithfully reproducing the physical behavior of the system,
- A set of explanatory functions with parameters included in the deterministic and/or the FE model.
- A Bayesian framework to obtain the correction parameters that properly account for model errors from an inaccurate model form or missing variables in the deterministic model.

As a result of this methodology we get a deterministic capacity (or demand) model, fitting a complex FEM, where the model error from the fitting process is explicit.

CHAPTER 3 - MODELLING PROCEDURE FOR EXPLICITING UNCERTAINTY SOURCES

3.1 Approach

Chapter 1 set the basis of the proposed approach. In this chapter, the proposed approach is derived. The role of the modeler can be synthetized in the following procedure (shown graphically in Figure 1)

Step 1. Selection of computational model candidates ($\boldsymbol{M}$) for predicting the behavior of the physical – mechanical process. Usually several options of different complexity arise.

Step 2. Selection of candidate load models ($\boldsymbol{L}$) Regarding the possible hazards affecting the structure under design, the designer must select the different possible hazards and the way in which each of them are going to be modeled. More than one candidate per hazard can arise.

Step 3. Name all the defined variables and parameters from steps 1 and 2.

Step 4. For all parameter and variables of Step 3, the modeler needs to assess the randomness (if any) and the correlation among variables that are going to be considered. In this step, the available information (expert judgment on the problem, databases and literature, and existing measurements) is used to build a probabilistic model of the input parameters, which is cast as a random vector $\boldsymbol{x}$**,** described by a joint probability density function (PDF) $f_x(\boldsymbol{x}, \Theta_f)$. If more than one PDF arises as result of this analysis there will be a set of candidate stochastic models ($\boldsymbol{\Omega}$). When the parameters are assumed statistically independent, this joint distribution is equivalently defined by the set of marginal distribution of all input parameters. If dependence exists, the copula formalism may be used (Phoon, et al., 2015).

Step 5 Construction of the explicit capacity surrogate models, as explained in 2.3.

Step 6. Construction of the explicit demand surrogate models, as explained in 2.3.

Step 7. Reliability analysis assessment.

Step 8 Sensitivity analysis performed based on the importance of the different variables. Four different importance measures are considered: (a) Variable importance, (b) stochastic model selection and correlation importance, (c) computational models selection importance and (d) load model selection importance.

By the end of the procedure, a detailed comparison between decisions made during the analysis is available in order to decide whether to improve the model. We don't necessary introduce a 'objective precision' model but we identify the main source of uncertainty, and can address it directly. Associating this model to the cost of addressing each of the errors committed we can have an optimization problem based on the economic cost of solving the problem.

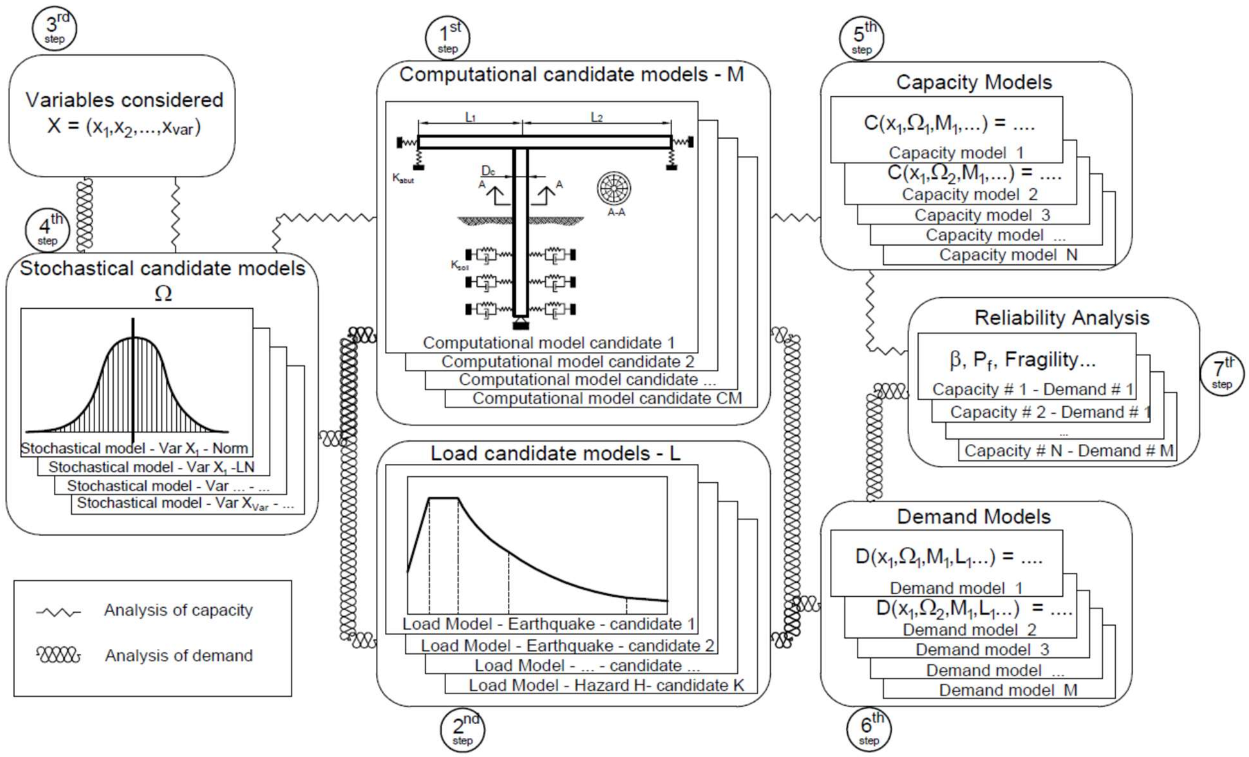

Figure 1 - Modelling framework

3.2 Inaccuracy in model form and volitional uncertainty

Steps 1 ,2 and 4 provide several options for the component model. Following the detailed definition of uncertainty types given in Murphy et al (2011) we can group the differences in the component model in two categories: inaccuracy in model form and volitional uncertainty.

Inaccuracy in model form is the error produced from assuming a simplification of the real (unknown) behavior of the physical phenomena (e.g. a linear expression is used when the actual relation is nonlinear). For example, for computing the resulting compression force in a reinforced concrete beam we can use the rectangular distribution developed by Whitney, or use a parabolic stress distribution.

Volitional uncertainty is associated with the decision made by the modeler in the component models behavior. For example, for modelling 3 seconds' gust wind speeds of extratropical winds several distributions are typically used: Gumbel, reverse Weibull, lognormal. If modelling this effect the engineer should define which one is going to apply, and there is no reason to believe that one is better than the other.

This approach made by the author is very important because is going to define how to handle these differences in behavior of the component models. Then the treatment we are giving to these two types of uncertainties is different.

Model form categorical variables are ranked since a more complex model captures a simpler model behavior (i.e. a quadratic interpolation is more powerful than a linear interpolation since the space of linear solutions are included in the quadratic space of solutions). Therefor we expect better results from more "complex" models. Hence, in this case, more complex models are assumed to be the correct answer, and comparing this solution with other approximate solution we can measure the error between model assumptions.

This situation is different in volitional uncertainty, since these model solutions are not ranked (i.e. when choosing between one behavior or the other one we do not know a priori which solution is the best). A good example of this occurs when selecting probabilistic functions $f_x(\boldsymbol{x}, \Theta_f)$ even though there are certain rules for choosing distributions as image space (for positive results lognormal distributions perform better than normal distributions, where the image space includes all the real numbers) this rules are usually non-conclusive on which model is better. etc. Then for the analysis of differences in volitional uncertainty between models a weighted average is going to be considered, where the weight, assigned by the modeler, considers the likelihood that the component model is the correct model. The variance of this weighted average is going to be the indicator of the volitional uncertainty. A more detailed description of this procedure is discussed in CHAPTER 4 - .

3.3 Definition of variables

Let us consider a physical system whose behavior can be represented by:

$$
\begin{aligned}
\mathbf{x} \in \boldsymbol{\Omega} \subset \mathbb{R}^{\mathrm{M}} \mapsto C_{\mathrm{k}} = \mathrm{y}_{\mathrm{k}}(\boldsymbol{x}, \boldsymbol{M}, \boldsymbol{\Omega}) \subset \mathbb{R} \\
\mathbf{x} \in \boldsymbol{\Omega} \subset \mathbb{R}^{\mathrm{R}} \mapsto \mathrm{D}_{\mathrm{k}} = y_{\mathrm{k}}(\boldsymbol{x}, \boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{L}) \subset \mathbb{R}
\end{aligned}
\tag{1}
$$

$$\mathrm{P_{f,k}}(\boldsymbol{x}, \boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{L}) = \mathrm{P}[\mathrm{C_k} - D_\mathrm{k} < 0]$$

Where:

$\boldsymbol{x} = (x_1, x_2, \ldots)$ gather the random variables representing input parameters for the model M (e.g. Soil parameters, material parameters, etc.)

y_k is the kth quantity of interest (QoI) in the analysis (e.g. load carrying capacity of a system, the maximum deflection, etc.).

$\boldsymbol{M} = (M_1, M_2, \ldots)$ gather the categorical variables that represent the different possible computational model that can predict the behavior of y_k (e.g. difference in the material considered, or the order of the approximation used, etc.)

$\boldsymbol{\Omega} = (\Omega_1, \Omega_2, \ldots)$ gather the categorical variables that represent the possible random space that model the randomness of the variables $\boldsymbol{x}$ (e.g. types of distributions, correlations, etc.).

$\mathbf{L} = (\mathrm{L}_1, L_2, \ldots)$gather the categorical variables that represent the different possible load model that can predict the hazard effect on y_k (e.g. response spectra, time history analysis, etc.)

$C_k$is the capacity of the structure in the direction k (shear, displacement, rotation, etc.).

D_k is the demand over the structure in the direction k (shear, displacement, rotation, etc.).

$\mathrm{P_{f,k}}$ is the probability of failure of the system in the k direction.

Recall that categorical variables can be: *model form categorical variables* or *volitional categorical variables*, following to the classification done before.

*3.3.1 Definition of **M**, computational model categorical variables*

As explained the set of computational models are those who per engineering judgment can predict the behavior of the system. The modeler should try to be as wide as possible in order to not leave out any physically possible computational model. The result of this analysis is a finite combination of possible computational models, which a priori could solve the physical problem. As seen in Figure 1 the selection of model behavior is going to impact the capacity and demand models. Typical different options on structural engineering are shown in Figure 2.

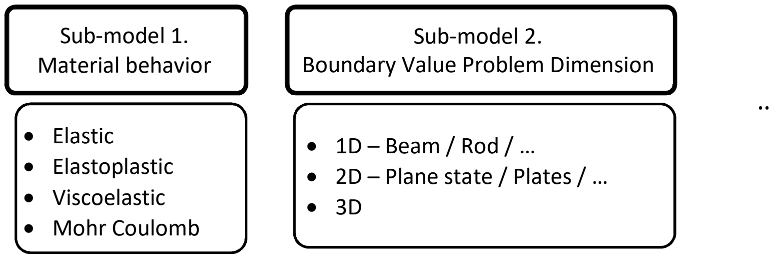

Figure 2 – Schematic representation of computational model (**M**) categorical variables for structural engineering

3.3.2 *Definition of L, Load model categorical variables*

The set of load models are those who per engineering judgment can predict the behavior of a certain hazard on the model. The result is again a finite combination of possible load models, which a priori could solve the physical problem. As seen in Figure 1 the selection of Load models is going to have impact only in the demand models. Typical different options on structural engineering are shown in Figure 3.

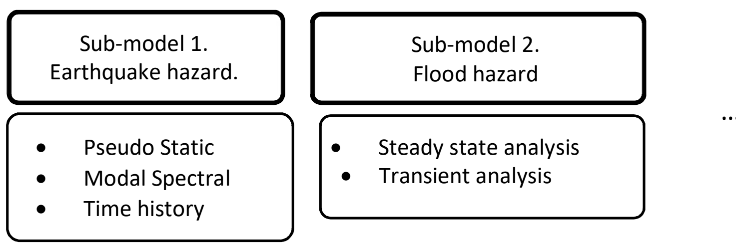

Figure 3 – Schematic representation of Load model (**L**) categorical variables selection

3.3.3 *Definition of Ω, random space categorical variables*

Defined all the possible computational model and load categorical variables we define the random distribution of the set of parameters needed and their possible correlation. The selection of this categorical variables can be represented in the flowchart shown in Figure 4. As in the previous section, a set of categorical random space variables is going to be the result of this step.

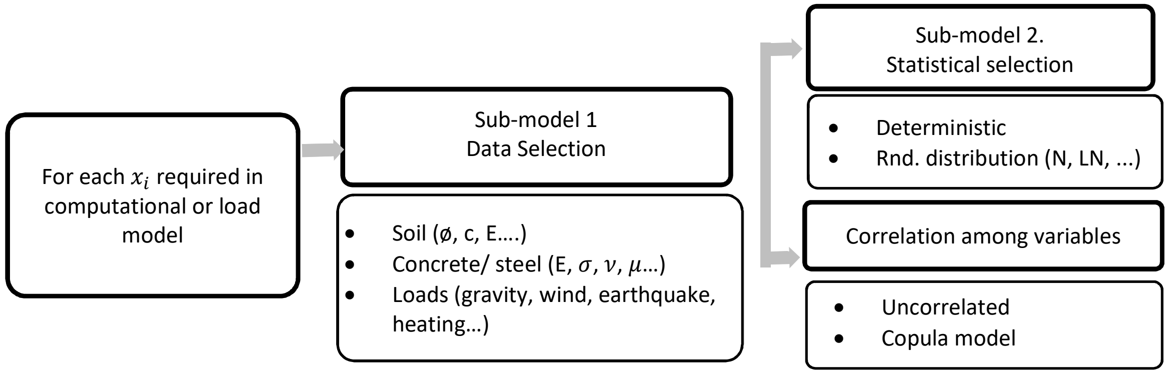

Figure 4 - Schematic representation of random space (Ω) categorical variables selection

3.4 Construction of behavioral models

3.4.1 Model form categorical variables

Defined all the possible categorical variables for the computational model, load model and the random space, by doing all the combinations among them we will have two sets of N 'behavioral models'. In Burnham et al (2002) this are called 'structural models', because the structure between them is different. Here the term structural is applied to the mechanic behavior of the construction, so we prefer to apply the term behavioral model. Since the behavior is different, the assumptions made are different.

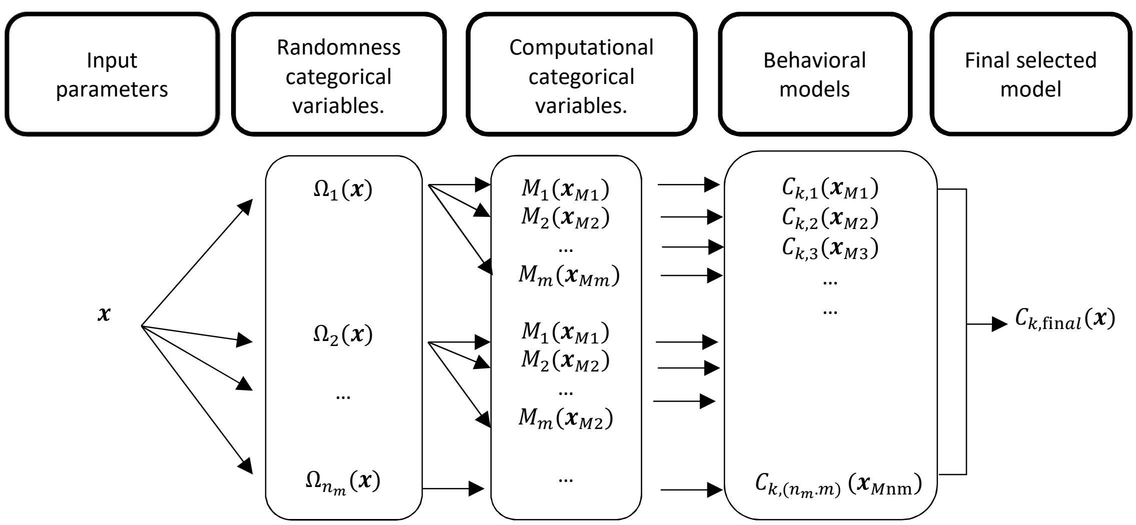

Figure 5 - Schematic representation of uncertainty propagation of the variables

One set for all the possible capacity models and other for all the demand models. Figure 5 show how the combination procedure is done for capacity models, when all the categorical variables are on model form behavior.

Note that $M_m(\boldsymbol{x}_{Mm})$ reflects that each model may require a different vector $\boldsymbol{x}_{Mm}$, the union of all $\boldsymbol{x}_{Mm}$ form $\boldsymbol{x}$. A similar construction needs to be done for the demand models D_k

The assessment of the selection among the possible options is going to be done in CHAPTER 4 -

3.4.2 Volitional uncertainty categorical variables

In the case of having a combination of model form categorical variables and volitional categorical variables, the display of possible combinations is different, since they are not ranked anymore, an array display like the one shown in Figure 6 is obtained.

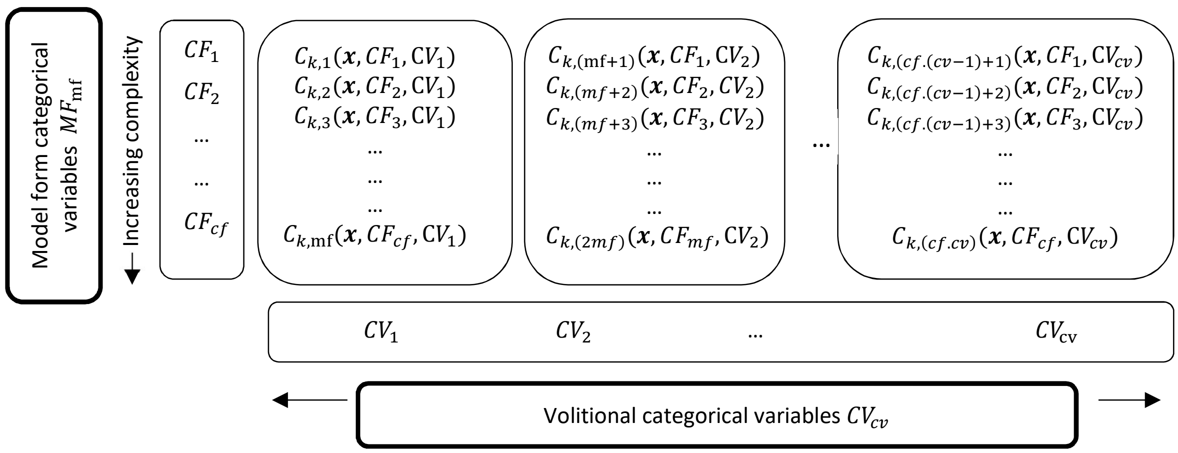

Figure 6 -Form model and volitional categorical variables

Where $CF_i, i = (1,2, \dots, cf)$, are those component models of the capacity that are classified as form model categorical variables, while $CV_i, i = (1,2, \dots, cv)$ are those components classified as volitional categorical variables. This component models for the capacity are combinations of computational model variables ($\boldsymbol{M}$) and stochastic model variables ($\boldsymbol{\Omega}$).

For the demand, a similar graph (not shown) can be obtained where we have DF and DV. In this case, the component models are a combination of computational model variables ($\boldsymbol{M}$), stochastic model variables ($\boldsymbol{\Omega}$), and load model variables ($\boldsymbol{L}$).

3.5 Matrix of categorical selection

As a result of previous selections, we get the required input for the analysis. Writing it in matrix form we get the matrix categorical selection. Each of the rows of the matrix corresponds to a different behavioral model. The first set of n columns are the set of inputs $\mathbf{x}$, with 1 if they are being considered in the respective behavioral model, or 0 otherwise. The following the set of columns displays the categorical variables, indicating for computation and demand which component model is considered, for both model from and volitional variables.

3.6 Surrogate model construction

Also referred as meta-model, or response surface, surrogate models refers to an explicit equation capable of representing the original model, i.e.:

$$y_{k,i}(\boldsymbol{x}) = y^{*}_{k,i}(\boldsymbol{x}) + \varepsilon \tag{2}$$

where $y_{k,i}(\underline{x})$ is the original model, $y^{*}_{k,i}(\underline{x})$ is the metamodel and ε is an approximation error. For the surrogate model construction, we follow the physics based functions (Gardoni, et al., 2002).

$$y^{*}_{k,i}(\boldsymbol{x}, \theta_k, \bar{\Sigma}) = \widehat{y_k}(\boldsymbol{x}) + \gamma_k(\boldsymbol{x}, \theta_k) + \sigma_k \varepsilon_k \text{ for } k = 1, \dots, q \tag{3}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots)$ denotes the set of unknown model parameters, $\bar{\Sigma}$ denotes the covariance matrix of the variables, $\widehat{y_k}(\mathbf{x})$ is the selected deterministic model, $\gamma_k(\mathbf{x}, \theta)$ is the correction term for the bias inherent in the deterministic model that is expressed as a function of the variables $\mathbf{x}$ and parameters $\boldsymbol{\theta}$, ε_k random variable with zero mean and unit variance, and σ_k represents the standard deviation of the model error. Note that for given $\boldsymbol{x}, \theta_k, \boldsymbol{\Sigma}$ and σ_k, we have $\mathrm{Var}[y_{k,Si}(\mathbf{x}, \theta_k, \boldsymbol{\Sigma})] = {\sigma_k}^2$ as the variance of the model.

The correction term can be expressed as:

$$\gamma_k(\mathbf{x}, \theta_k) = \sum_{i=1}^{p_k} \theta_{ki} \; h_{ki}(x) \text{ for } k = 1, \dots, q \tag{4}$$

where $h_{ki}(\mathbf{x})$ are the explanatory functions. These functions should be selected to enhance the predictive capacity of the deterministic model $\widehat{y_k}(\mathbf{x})$. It is appropriate to select terms that are thought to be missing in $\widehat{y_k}(\mathbf{x})$. Ideally, rules of mechanics should be used in formulating the explanatory functions. However, in many cases reliance on intuition is necessary. It is also desirable that $h_{ki}(\mathbf{x})$ have the same dimension as $\widehat{y_k}(\mathbf{x})$ so that θ_{ki} are dimensionless. It is recommended to start the model assessment process with a comprehensive candidate form of $\gamma_k(\mathbf{x}, \theta_k)$ and then simplify it by deleting unimportant terms or combining them when highly correlated (Gardoni, et al., 2002).

For fitting the functions described above we need to get enough samples of the model $y_{k,i}(\boldsymbol{x})$ in order to approximate $y^*_{k,i}(\boldsymbol{x})$. Generally, the computational model for solving $y_{k,i}(\boldsymbol{x})$ is a complex FEM, therefore we need to achieve a sampling technique capable of representing the random space with the least samples possible in order to get acceptable times for running $y_{k,i}(\boldsymbol{x})$. For this purpose, Latin Hypercube sampling is adopted (Iman, 2008). An enhanced methodology is applied that besides modelling the complete random space, ensures the designed correlation (Vořechovský, et al., 2009).

CHAPTER 4 - ASSESING THE SOURCE OF UNCERTAINTY OVER THE SYSTEM RELIABILITY

In the procedure presented in CHAPTER 3 - we have two sets of easily evaluable models (N_C capacity models and N_D demand models), and the model error of the approximation of the surrogate model to the FEM, if applicable. We also have the different categorical variables grouped per sub-model selection.

The next step is to perform a reliability analysis over the combination of this models to evaluate the probability of failure (P_f) and the reliability index (β).

4.1 Comparison of volitional categorical variables

When considering volitional categorical variables we have two sources of uncertainty, the first one associated with the variability of the input $\boldsymbol{x}$, the second one associated with the uncertainty related with which behavioral model better represents better the physical problem.

The existence of several behavioral models implies the existence of several limit state surfaces. To compare the different hypersurfaces we must collect them in the same space, since different volitional models we can have different sets of inputs $\boldsymbol{x}$. To compare the different limit state functions, we compare the hypersurface in the space of greater dimension, assigning a value of 0 to the subset of $\boldsymbol{x}$ not considered by the reduced dimension models. By the definition of volitional variables we are not certain about which of them is the correct one, then the definition of failure is not trivial, and we are going to have a certain failure subspace and a likely failure subspace, as shown in Figure 7.

The limit state function in this case can be defined as:

$$G_k(\boldsymbol{u}, V_{\mathrm{k,j}}) = {G_k}^V(\boldsymbol{u}|V_{\mathrm{k,j}}).P(V_{\mathrm{k,j}}), \qquad \mathrm{j} = (1 \dots \mathrm{bv}) \tag{5}$$

where $V_{\mathrm{k,j}}$is the volitional selection, and $P(V_{\mathrm{k,j}})$ is the assigned likelihood of the categorical variable $V_{\mathrm{k,j}}$. The sub-index k represents the direction of the analyzed failure type (shear, flexure, displacement, …), and will be omitted from now on to simplify the mathematical derivation.

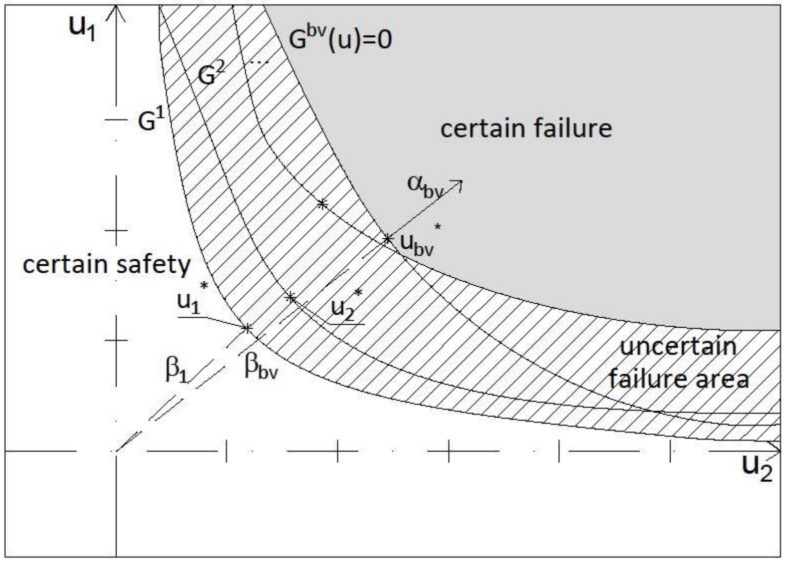

Figure 7 - Limits states functions in standard normal space

We analyze first the uncertainty that the variability in the input random variables has on the probability of failure. For each one of the limit states functions, following the procedure developed in Der Kiureghian (2005), we first do the transformation of the random space $\boldsymbol{x}$ to the standard normal random space $\boldsymbol{u}$. Then $g^V(\boldsymbol{x}|V_{\mathrm{j}})$ the limit state function in the x space, is $G^V(\boldsymbol{u}|V_{\mathrm{j}}) \equiv g^V(T^{-1}(\boldsymbol{u})|V_{\mathrm{j}})$ in the $\boldsymbol{u}$ space, V_{j} are the *bv* different volitional component models. Recall the definition of the design point:

$$u^{V*} = \arg min\{\| \mathbf{u} \| \, |G^V(\boldsymbol{u}|V_{\mathrm{j}}) = 0\} \tag{6}$$

We linearize the limit state function at u*, the design point.

$$\begin{gathered} G^V(\boldsymbol{u}|V_{\mathrm{j}}) \approx G_1{}^V(\boldsymbol{u}|V_{\mathrm{j}}) = G^V(\boldsymbol{u}^*|V_{\mathrm{j}}) + \nabla G(\boldsymbol{u}^*|V_{\mathrm{j}}).(\boldsymbol{u}-\boldsymbol{u}^*) \;\; \mathrm{j} = (1 \ldots \mathrm{bv}) \\ G_1{}^V(\boldsymbol{u}|V_{\mathrm{j}}) = \| \nabla G(\boldsymbol{u}^*|V_{\mathrm{j}}) \| \, (\beta_j - \boldsymbol{\alpha u}) \;\; \mathrm{j} = (1 \ldots \mathrm{bv}) \\ where, \qquad \boldsymbol{\alpha} = \frac{-\nabla G(\boldsymbol{u}^*|V_{\mathrm{j}})}{\| \nabla G(\boldsymbol{u}^*|V_{\mathrm{j}}) \|} \quad ; \quad \beta = \boldsymbol{\alpha}.\boldsymbol{u}^* \quad ; \quad P_f = \Phi(-\beta) \end{gathered} \tag{7}$$

Computing the expectations of G:

$$\begin{gathered} E_U[G_k(\boldsymbol{u}, V_{\mathrm{k,j}})] = E_U[G^V(\boldsymbol{u}|V_{\mathrm{j}})].\; P(V_{\mathrm{j}}) \approx \| \nabla G(\boldsymbol{u}^*|V_{\mathrm{j}}) \| \, \beta_j . P(V_{\mathrm{j}}) \\ Var_U[G_k(\boldsymbol{u}, V_{\mathrm{k,j}})] = Var_U[G^V(\boldsymbol{u}|V_{\mathrm{j}})].\; P^2(V_{\mathrm{j}}) \approx \| \nabla G(\boldsymbol{u}^*|V_{\mathrm{j}}) \|^2 . P^2(V_{\mathrm{j}}) \\ \beta_j = \beta(V_{\mathrm{j}}) = \frac{E_U[G_k(\boldsymbol{u}, V_{\mathrm{k,j}})]}{\sqrt{Var_U[G_k(\boldsymbol{u}, V_{\mathrm{k,j}})]}}, \qquad \mathrm{j} = (1 \ldots \mathrm{bv}) \end{gathered} \tag{8}$$

Having this set of limit state functions, we define the most likely limit state hypersurface as the weighted average over the set of limit state surfaces. Correspondingly we define the mean reliability index $\bar{\beta}$ as the weighted average of the set of j reliability indexes:

$$E_V[\beta(V_{\mathrm{j}})] = \sum_{j=1}^{bv} \beta_j . P(V_{\mathrm{j}}) = \bar{\beta} \tag{9}$$

We compute as well the variation over this weighted average.

$$Var_V[\beta(V_{\mathrm{j}})] = \left[\sum_{j=1}^{bv} P(V_{\mathrm{j}}).\sqrt{Var_U[G_k(\boldsymbol{u}, V_{\mathrm{k,j}})] + (\beta_j - \bar{\beta})^2}\right]^2 \tag{10}$$

To derive measures of importance for the basic random variables, consider linearizing the transformation $\boldsymbol{u} = T(\boldsymbol{x})$ at the design point $\boldsymbol{u}^*$:

$$\begin{gathered} \boldsymbol{u} \cong \boldsymbol{u}^* + J_{u,x}(\boldsymbol{x} - \boldsymbol{x}^*) \\ \boldsymbol{u} = \boldsymbol{u}^* + J_{u,x}(\widehat{\boldsymbol{x}} - \boldsymbol{x}^*) \end{gathered} \tag{11}$$

Where x^* is the inverse of the transformation of $\boldsymbol{u}^*$ $(i.e.\ \boldsymbol{x}^* = T^{-1}(\boldsymbol{u}^*))$. $J_{u,x}$ is the Jacobian of the transformation from the random space of $\boldsymbol{x}$, to the standard normal space $\boldsymbol{u}$. Note $\widehat{\boldsymbol{x}}$ is slightly different from $\boldsymbol{x}$ because $\widehat{\boldsymbol{x}}$ is a linear function of $\boldsymbol{u}$, it must have the joint normal distribution. Its covariance matrix is:

$$\Sigma \cong {J_{u,x}}^{-1}({J_{u,x}}^{-1})^T \tag{12}$$

Now from Eq. (11) we can derive the variance on $G_k(\boldsymbol{u}, V_{\mathrm{k,j}})$ as:

$$Var_U[G_1(\boldsymbol{u}, V_{\mathrm{j}})] = \parallel \nabla G(\boldsymbol{u}^* | V_{\mathrm{j}}) \parallel^2 . (\parallel \boldsymbol{\alpha}.J_{u,x}.\widehat{D} \parallel^2 + \boldsymbol{\alpha}.J_{u,x}.(\widehat{\Sigma} - \widehat{D}.\widehat{D}).{J_{u,x}}^T.\boldsymbol{\alpha^T}) \tag{13}$$

Where $\widehat{D} = diag[\widehat{\boldsymbol{\sigma}}_{\boldsymbol{\iota}}]$ is the diagonal matrix of standard deviations of $\widehat{\boldsymbol{x}}$. The first term in the above expression contains the contributions to the variance of $G_1(\boldsymbol{u}, V_{\mathrm{j}})$ arising from the individual variances of the elements of $\hat{x}$, whereas the second term represents the contributions arising from the covariances of pairs of the random variables. Hence, the elements of the vector $\parallel \nabla G(\boldsymbol{u}^* | V_{\mathrm{j}}) \parallel . \left[\overline{\boldsymbol{\alpha}.J_{u,x}.\widehat{D}}\right]$ can be considered to provide relative measures of importance of the elements of $\widehat{\boldsymbol{x}}$, or approximately of $\boldsymbol{x}$.

Eq. (10) can now be written as:

$$Var_V\left[\beta(V_j)\right] = \left[\sum_{j=1}^{bv} P(V_j).\sqrt{\| \nabla G(\boldsymbol{u}^*|V_j) \|^2.(\| \alpha.J_{u,x}.\widehat{D} \|^2) + (\beta_j - \bar{\beta})^2}\right]^2$$

$$\sigma_V[\bar{\beta}] = \sum_{j=1}^{bv} P(V_j).\gamma_{vol,j} \tag{14}$$

$$\gamma_{vol,j} = \frac{\left[\| \nabla G(u^*|V_j) \|.\left[\overline{\alpha.J_{u,x}.\widehat{D}}\right] \quad |\beta_j - \bar{\beta}|\right]}{\| \left[\| \nabla G(u^*|V_j) \|.\left[\overline{\alpha.J_{u,x}.\widehat{D}}\right] \quad |\beta_j - \bar{\beta}|\right] \|}$$

We get a modified importance vector, where we just append a vector of the absolute differences between β_j the reliability index of the j limit state functions and the mean $\bar{\beta}$. Note that in the case of not having any volitional categorical variable, we get the same expression as the importance vector γ defined in the referenced paper (Der Kiureghian, 2005).

With this modified importance vector we are capable of comparing the importance of both the random variables representing input parameters for the model $\boldsymbol{M}$ and volitional variables representing different model behavior. As this importance measures are taken with respect to the variation that these uncertainties have in the probability of failure, it is a fair comparison for both measures, allowing the modeler to decide whether to consider the volitional variables have any important effect or whether the main uncertainties are in the input parameters $\boldsymbol{x}$.

4.2 Comparison of form model categorical variables

If we arrange the set of reliability index (β) in an array where columns are corresponding to volitional categorical variables and the rows are the form model categorical variable, we can get a set of ranked reliability indexes by expressing the more likely reliability index as representative of the model form accuracy, as shown in Figure 8.

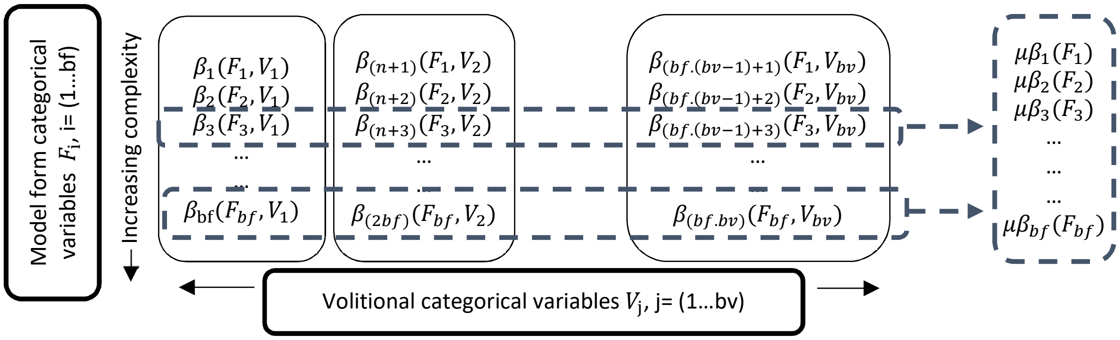

Figure 8 - Comparison of form model categorical variables

Where $F_i\ for\ i = (1, 2, \dots, bf)$ are those component models of the system that are classified as form model categorical variables, while $F_j\ for\ j = (1, 2, \dots, bv)$ are those components classified as volitional categorical variables. This component models are combinations of computational model variables $(\boldsymbol{M})$, stochastic model variables $(\boldsymbol{\Omega})$ and load categorical variables $(\boldsymbol{L})$. Then bf is equal to the product of form model categorical variables corresponding to $\boldsymbol{M}$, $\boldsymbol{\Omega}$ and $\boldsymbol{L}$ (i.e. $\boldsymbol{bf} = (\boldsymbol{M_f}.\boldsymbol{\Omega_f}.\boldsymbol{L_f})$). The same for **bv** as is the product of volitional categorical variables corresponding to $\boldsymbol{M}$, $\boldsymbol{\Omega}$ and $\boldsymbol{L}$ (i.e. $\boldsymbol{bv} = (\boldsymbol{M_v}.\boldsymbol{\Omega_v}.\boldsymbol{L_v})$).

The set of $\mu\beta$ are representative of the uncertainty in the model formulations, since the bigger the uncertainty in our input values, the higher the probability of failure. The more precise model is going to be set the corresponding to $\mu\beta_{bf}$, and the error on the other models is going to be computed with respect to this model. To decide which of the models is more appropriate we couple the uncertainty with the model complexity.

Several definitions for model complexity (O) are available where different properties of the model are being evaluated. In this case, we focus on the time required by the model to run and the number of variables.

$$\mathrm{O}(bf) = \text{time required (hr)} + Number\ of\ variables\ for\ computing\ (\mu\beta_{bf}) \qquad (15)$$

By plotting the pairs $(\mu\ \ , O)$ we build the complexity uncertainty graph (Figure 9). The set of best models are going to form a Pareto front, since by the definition of this kind of front, this points are going to show the least complexity for a given uncertainty, or the least uncertainty for a given complexity. The models marked with an (*) are not going to be solution of the problem, but is a decision of the modeler to choose, among the points over the Pareto front, the model to be considered as the best.

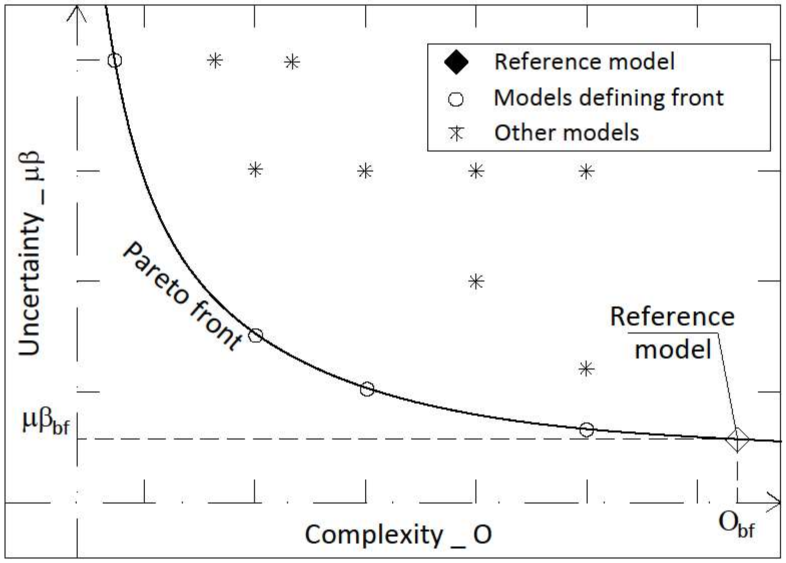

Figure 9 - Comparison of form model variables

CHAPTER 5 - EXAMPLE CANTILEVER BEAM ON CLAY FOUNDATION

Let us consider the cantilever column shown in Figure 10, with a shallow foundation over a clay deposit. The direction of failure is assigned to be the deflection of the tip of the column. The following sources of uncertainty are considered:

- Uncertainty in Soil elasticity parameters for Sand Material (E_{clay}, μ_{clay})
- Uncertainty in crack development over the cantilever column, represented by its effective modulus of inertia (I_{ef})
- Uncertainty in the lumped load, with known mean and standard deviation, but unknown distribution. Two options are considered for the distribution of P: Lognormal or Type I maximum.

For simplicity all materials are considered to be perfectly elastic and the capacity is deterministic and set to L/120, as is typically required in design codes.

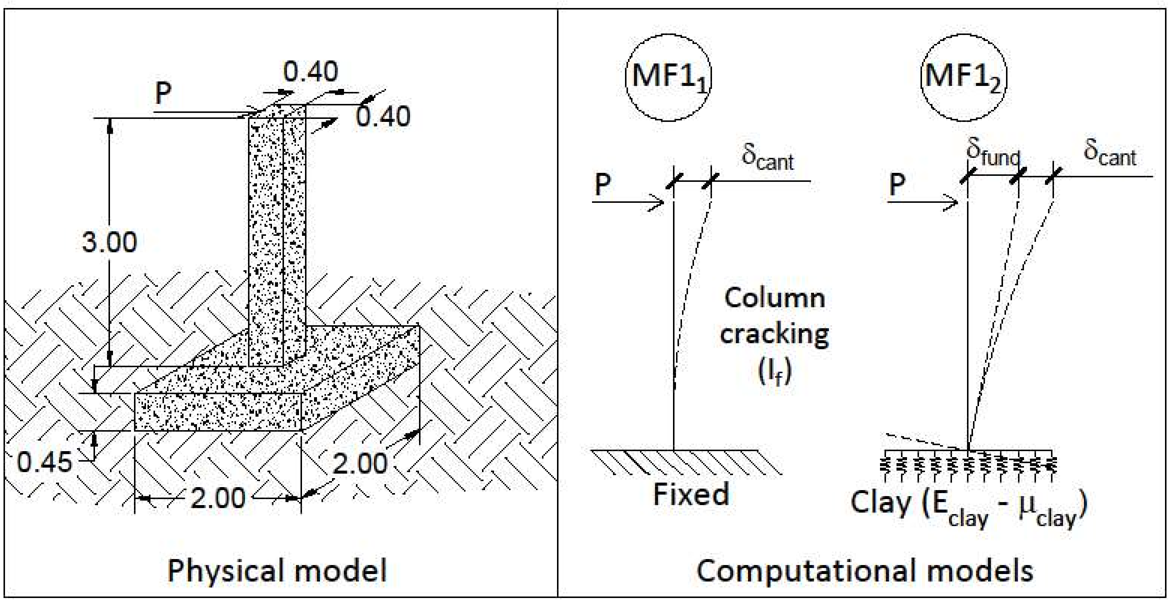

Figure 10 - Example model

5.1 Definition of variables

5.1.1 Definition of M, computational model categorical variables

As stated in the problem definition all materials are considered to behave linearly elastic. Then the possible solutions considered are:

$$MF1_1 \longrightarrow C(\boldsymbol{x}) = \delta_{cant}(\boldsymbol{x})$$
$$MF1_2 \longrightarrow C(\boldsymbol{x}) = \delta_{cant}(\boldsymbol{x}) + \delta_{found}(\boldsymbol{x}) \tag{16}$$

We have defined one categorical variable, called it MF1, with options $MF1_1$ and $MF1_2$, as shown in Figure 10 - Example model. This categorical variable is a form model categorical variable, since $MF1_2$ is a more precise solution than $MF1_1$. In this example other suggestions could be made, but for the sake of simplicity we are keeping only one alternative.

The calculation of the deflection of the tip of is equal to that of the elastic behavior of the column for $MF1_1$ and the previous result plus the rotation from the elastic behavior of the soil for $MF1_2$. The elastic solution of the deflection of the cantilever is:

$$\delta_{cant} = \frac{P_t L_c^3}{3 E_b I_{ef}} \tag{17}$$

Where: P_t is the load applied in the tip of the cantilever, L_c the length of the cantilever, $E_b I_{ef}$ the effective modulus of elasticity of the beam (in this case we considered it reduced by cracking, then $E_b I_{ef} = E_b I_g f$, with I_g the modulus of inertia of the gross section and $f \in (0,0.6]$ the cracking variable.

Tip deflection due to base rotation is computed following Bowles (1988).

$$\delta_{found} = \frac{1 - \mu_{soil}^2}{E_{soil}} \frac{P_t\, L_c^2}{b_f^2\, L_f} I_\theta \tag{18}$$

Where E_{soil} and μ_{soil} are the modulus of elasticity and Poisson coefficient of the soil, respectively; L_f and b_f the dimensions of the base and I_θ the influence values, that incorporates the stiffness ratio between the foundation and the soil, in this case is for square base L_f/B_f =1, $I_\theta = 4.17$.

5.1.2 Definition of L, computational model categorical variables

The definition of the load model is straight forward, with the load P applied in the tip. No other alternatives are considered.

5.1.3 Definition of Ω, random space categorical variables

Defined the parameters and variables in 5.1.1 and 5.1.2 we must first define which ones are deterministic (a parameter from now on) and which are considered to be variable. Then, for the variables, we need to define the type of distribution, and finally to assess if there exists any correlation between them.

In this example are considered variables the effective flexural rigidity ($x_1 = E_b I_{ef}$), the clay elasticity modulus ($x_2 = E_{clay}$), the clay poisson coefficient ($x_3 = \mu_{clay}$) and the intensity of the load P ($x_4 = P$).

There are known bounds for the possible values of (x_1, x_2, x_3). The least informative distribution with known bounds is the uniform distribution, then we assign this distribution to these variables.

For x_4, as stated in the problem, there is considered to be certainty about the mean and the standard deviation. About the distribution there is no such certainty, then following the procedure we define the distribution of P, $\Omega\mathrm{V}1$, as a random space categorical variable. Two options are available: $\Omega\mathrm{V}1_1$ (lognormal distribution of P) and $\Omega\mathrm{V}1_2$ (Type I max distribution for P). In this case as not enough information is available to define which of the distributions is the correct one, this variable is a volitional categorical variable. The variables distributions are shown in Table 1.

Table 1 - Random space definition

$\boldsymbol{x}$ - *Input variables*	Distribution	Mean (μ)	Standard dev. (σ)	Coefficient of var. ($c.o.v.$)	Correlations (ρ_{ij})
x_1 – *Column flexural rigidity [kN.m^2]*	Uniform	23.8	5.9	0.25	
x_2 - *Clay elasticity modulus [MPa]*	Uniform	51	28.3	0.55	
x_3 - *Clay Poisson coef. [adim]*	Uniform	0.45	0.029	0.064	$Identity_{4x4}$
x_4 - *Load Intensity. [kN] -* $\Omega 1_1$	Lognormal	98	9.8	0.1	
X_4 - *Load distribution* $\Omega 1_2$*. [kN]*	Type 1 max	98	9.8	0.1	

5.2 Construction of behavioral models

For the capacity model the maximum deflection is set in the problem statement as $L_C/120$. For the demand model we show the possibilities in Figure 11.

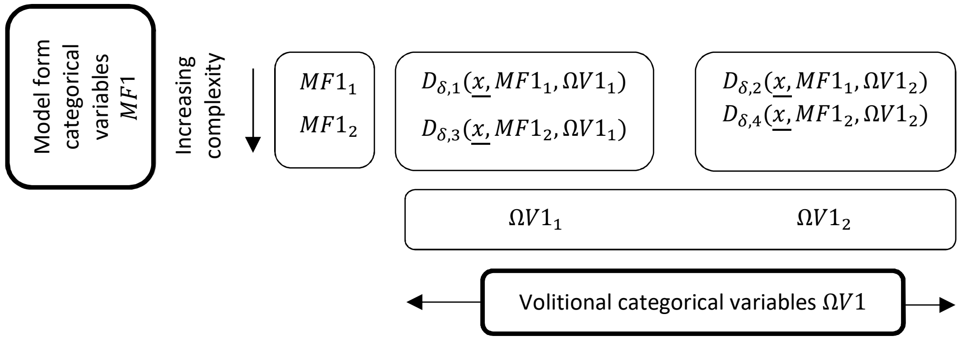

Figure 11 - Schematic representation of demand model construction

Once the capacity and demand models are defined, we can compose them to create the behavioral models. As we have only one capacity model and 4 demand models, we get 4 behavioral models (B_i), each one of them associated to one limit state function ($\mathrm{G_i}$) as the difference of the capacity model and one of the demand models, i.e.:

$$\mathrm{B}_i \longrightarrow \mathrm{G_i}(\mathbf{x}) = \mathrm{C}(\mathbf{x}) - \mathrm{D_{\delta,i}}(\mathbf{x}) \tag{19}$$

5.2.1 *Matrix of categorical selection*

In order to visualize the behavioral model selection, we show in Table 2 the matrix of categorical selection, as defined in 3.5. Each of the rows of the matrix corresponds to a different behavioral model. The first set of columns are the set of inputs $\mathbf{x}$, with 1 if they are being considered in the respective behavioral model, or 0 otherwise. The following the set of columns displays the categorical variables, indicating for computation and demand, which component model is considered.

Table 2 - Matrix of categorical selection

Behavioral models	*Continuous variables*					*Discrete Variables*		
						Cap. model	Demand model	
	x_1	x_2	x_3	x_4	$\mathrm{R}_{x,x}$	C	ΩV1	MF1
$B_1 \longrightarrow G_1(\boldsymbol{x}) = C - D_{\delta,1}$	1	0	0	1	$Identity_{4x4}$	1	1	1
$B_2 \longrightarrow G_2(\boldsymbol{x}) = C - D_{\delta,2}$	1	0	0	1		1	2	1
$B_3 \longrightarrow G_3(\boldsymbol{x}) = C - D_{\delta,3}$	1	1	1	1		1	1	2
$B_4 \longrightarrow G_4(\boldsymbol{x}) = C - D_{\delta,4}$	1	1	1	1		1	2	2

5.2.2 Surrogate model construction

In this case as we have already explicit equations there is no necessity of a surrogate model.

5.3 Assessing sources of uncertainty over reliability of the system

The solution of the four behavioral models is computed in FERUM, the importance measures are obtained from a first order reliability method (FORM) analysis, while for the probability of failure from a subset Monte Carlo (MC) simulation. The results are shown in Table 3.

Table 3 - Analysis results for probability of failure

Behavioral models	*Importance measures*				β	P_f
	x_1	x_2	x_3	x_4		
B_1	-0.32	-	-	0.94	**7.62**	**1.23 e-14**
B_2	-0.38	-	-	0.92	**5.38**	**3.62 e-8**
B_3	-0.43	-0.84	-0.09	0.33	**1.85**	**0.03146**
B_4	-0.43	-0.82	-0.09	0.95	**1.86**	**0.03233**

As we can see, the probability of failure considering only the deflection of the cantilever (B_1 and B_2) almost zero compared to the model considering the soil deflection (B_3 and B_4).

5.3.1 Comparison of form model categorical variables

Following the procedure defined in 4.2, we compute the average of same form model categorical variables. The results are shown in Figure 12.

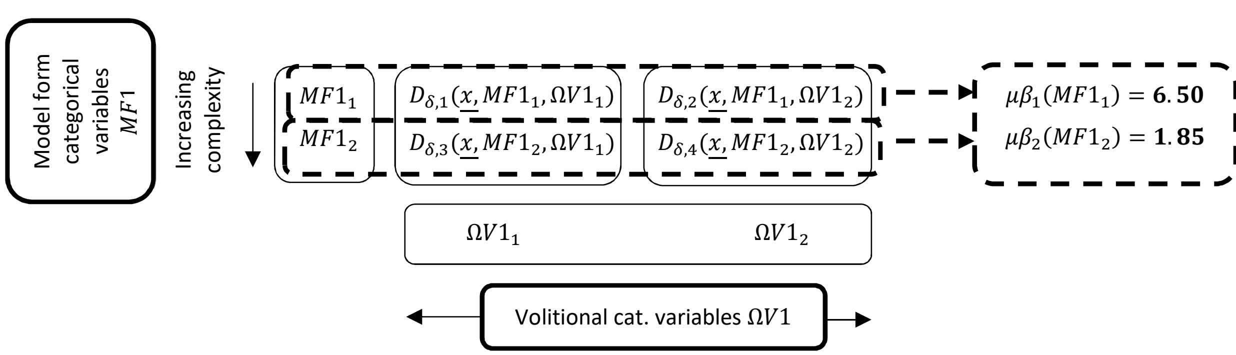

Figure 12 - Example comparison of form model categories

As in this simple case we only have two points, it is considered useless to show a Pareto front. The complexity of the models is 3 for $MF1_1$ and 5 for $MF1_2$.

5.3.2 Comparison of volitional categorical variables

From the results of Table 3 and Figure 12, we can infer that the approximation of not considering the soil behavior leads to highly unsafe results. Therefore, we continue the analysis of the behavioral models B_3 and B_4. Following the procedure defined in 4.1, we compute the modified importance measures, considering that the two volitional categorical variables ($\Omega V1_1$ and $\Omega V1_2$) are equally likely. This consideration, is the least informative assumption for the likelihood selection. The result for the reliability index and the importance measures are shown below, and in Figure 13.

$$\bar{\beta} = \sum_{j=1}^{2} \beta_j . P(V_j) = 1.85 \, . \frac{1}{2} + 1.86 \, . \frac{1}{2} = 1.855$$

$$\gamma_{vol,j} = \frac{\left[\parallel \nabla G(u^* | V_j) \parallel . \left[\overline{\alpha . J_{u,x} . \widehat{D}} \right] \quad |\beta_j - \bar{\beta}| \right]}{\parallel \left[\parallel \nabla G(u^* | V_j) \parallel . \left[\overline{\alpha . J_{u,x} . \widehat{D}} \right] \quad |\beta_j - \bar{\beta}| \right] \parallel} \tag{20}$$

$$\gamma_{vol} = [(-0.32 \quad -0.63 \quad -0.066 \quad 0.26) \quad (0.01)]$$

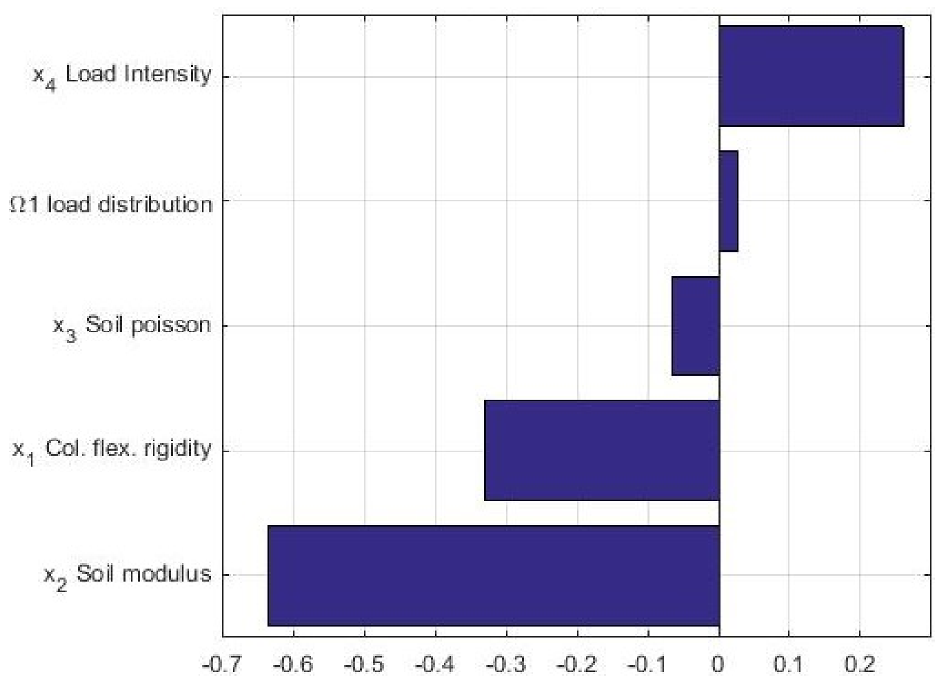

Figure 13 - Bar diagram of Importance measures

5.4 Result analysis for the example 1

The comparison of the form model categorical variables shows clearly that analyzing this system as a cantilever beam leads to very inaccurate results. The differences in the probability of failure are 10 times smaller than those obtained by a simple modelling of elastic behavior of the soil. Under these results, we should conclude that any model should consider soil behavior, since the deflection produced by the rigid rotation has more importance than the one associate to the fissuration of the column.

By the analysis of the volitional variables on the other hand, we see that the definition of the distribution of the load has a very small importance. Hence, to fix the distribution does not have great effects and will simplify following analysis. The decision of which distribution to choose is a modeler decision, but a usual criterion is to make the decision laying to the one in the conservative side.

Also, we learn from the procedure that to reduce the global uncertainty, we should focus our attention to get more data from the soil behavior and from the properties of the cracked section of the column, since these two variables are the ones with the highest importance (-0.6 and -0.3, respectively). Doing any of them separately is not going to improve the results greatly since, as stated in the beginning the final uncertainty is function of the greater source of uncertainty.

CHAPTER 6 - CONCLUSION FOR THE PROCEDURE PROPOSED

A simple methodology for quantifying uncertainty, and clarifying uncertainty sources was developed and showed in practice through a simple example. Several problems have been addressed by this procedure.

First, by proposing the incorporation of categorical variables in a reliability analysis we are able to address multi-model uncertainty. This incorporation allows to address the uncertainty coming from different candidate models, which was not considered in previous models applied to civil engineering. Improving this way traditional reliability analysis, where only uncertainty from input random variable $\boldsymbol{x}$ was considered.

Aligned with this enhancement, the distinction between model form categorical variables, where different variables are ranked and one candidate is better than the other, and volitional uncertainty categorical variables, where all candidates have some likelihood of being correct, allows the modeler to include properly, different types of candidates in consideration when modelling.

The consideration of categorical variables, and of different types, demand an improvement in the way in which the importance measures is considered. This is achieved by the incorporation of the enhanced importance, which fairly compares the importance that both, random variables and categorical variables, have in the probability of failure.

All these incorporations are done considering simple tools as: (a) a regular math software, (b) a finite element software and (c) open coded reliability software. All tools that are frequent in professional and research practice.

The system uncertainty is a function of the biggest source of uncertainty. Then the usefulness of this procedure is to address this source. Is later role of the modeler to assign the resources to reduce the system uncertainty. Basically, two options are possible to achieve this goal. The first one is to reduce the uncertainty of the inputs (i.e. improve the stochastic models $f_x(\boldsymbol{x}, \Theta_f)$), reducing its variance by gathering more data. The second option is to refine the models used to predict the physical behavior (i.e. improve $g_i(\boldsymbol{x}, \Theta_g)$) with more complex formulations. Attention should be paid to the latter option since there is a limit to model refinement dictated by the uncertainty in the inputs and the model formulations.

The main drawback of the methodology is that several computations are required to address the system inaccuracy. But programed efficiently, the programmer must let modern computers do most of the work.

REFERENCES

Bowles J. E. Foundation Analysis and Design, 4th Edition [Book]. - [s.l.] : McGraw-Hill, 1988.

Burnham K. and Anderson D. Model Selection and Multimodel Inference [Book]. - New York : Springer, 2002. - 2nd.

Der Kiureghian A. and Ditlevsen O. Aleatory or epistemic? Does it matter? [Journal] // Struct. Saf.. - 2009. - Vol. 31(2). - pp. 105–112.

Der Kiureghian A. and Ke J.B, Finite-element based reliability analysis of frame structures. [Journal] // In Proc. 4th Int. Conference on Structural Safety and Reliability. - 1985. - pp. 395-404.

Der Kiureghian A. First-and second-order reliability methods [Journal] // Engineering design reliability handbook. - 2005. - Vols. 14-11..

Der Kiureghian A., Haukaas T. and Fujimura K. Structural reliability software at the University of California, Berkeley [Journal] // Structural Safety. - 2006. - pp. 44-67.

Field Jr R. V. and Grigoriu M. Model selection in applied science and engineering: A decision-theoretic approach [Journal] // Journal of Engineering Mechanics. - 2007. - Vol. 133(7). - pp. 780-791.

Gardoni P., Der Kiureghian A. and Mosalam K. Probabilistic Capacity Models and Fragility Estimates for Reinforced Concrete Columns based on Experimental Observations [Journal] // J. Eng. Mech.. - 2002. - Vol. 128 (10). - pp. 1024-1038.

Gardoni P., Mosalam K. M. and Der Kiureghian A. Probabilistic seismic demand models and fragility estimates for RC bridges [Journal] // Journal of Earthquake Engineering. - 2003. - Vol. 7(spec01). - pp. 79-106.

Haukaas T. and Der Kiureghian A. FERUM [Online] // http://www.ce.berkeley.edu/projects/ferum/. - 2005.

Huang Q., Gardoni P. and Hurlebaus S. Probabilistic seismic demand models and fragility estimates for reinforced concrete highway bridges with one single-column bent [Journal] // Journal of engineering mechanics. - 2010. - Vol. 136(11). - pp. 1340-1353.

Iman R. L. Latin hypercube sampling. [Book]. - New York : Wiley, 2008.

Marelli S., Schobi R. and Sudret B. Structural Reliability, Report UQLab-V0.9-107 [Book Section] // UQLab user manual. - Zurich : Chair of Risk, Safety & Uncertainty Quantification ETH Zurich, 2015.

Murphy C., Gardoni P. and Harris C. E. Classification and moral evaluation of uncertainties in engineering modeling [Journal] // Science and engineering ethics. - 2011. - Vol. 17(3). - pp. 553-570.

Patelli E. [et al.] General purpose software for efficient uncertainty management of large finite element models [Journal] // Finite elements in analysis and design. - 2012. - Vol. 51. - pp. 31-48.

Pellissetti M. F. and Schuëller G. I. On general purpose software in structural reliability–An overview [Journal] // Structural Safety. - 2006. - Vol. 28(1). - pp. 3-16..

Phoon K.K. and Ching J. Y. Risk and Reliability in Geotechnical Engineering [Book]. - Florida : CRC Press, 2015.

Stefanou G The stochastic finite element method: Past, present and future [Journal] // Comput. Methods Appl. Mech Eng.. - 2009. - Vol. 198. - pp. 1031-1051.

Sudret B. and Der Kiureghian A. Comparison of finite element reliability methods [Journal] // Probabilistic Engineering Mechanics. - 2002. - Vol. 17(4). - pp. 337-348..

Sun B. and Gardoni P. A model selection for Hierarchy Model System in structural engineering [Journal]. - (unpublished).

Tabandeh A. and Gardoni P. Empirical Bayes Approach for Developing Hierarchical Probabilistic Predictive Models and Its Application to the Seismic Reliability Analysis of FRP-Retrofitted RC Bridges [Journal] // ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering. - 2015. - Vol. 1. - p. 04015002.

Tabandeh A. and Gardoni P. Probabilistic capacity models and fragility estimates for RC columns retrofitted with FRP composites. [Journal] // Engineering Structures. - 2014. - Vol. 74. - pp. 13-22.

Vořechovský M. and Novák D. Correlation control in small-sample Monte Carlo type simulations I: A simulated annealing approach. [Journal] // Probabilistic Engineering Mechanics. - 2009. - pp. 24(3). 452-462.